16

Altmetrics Gaming: Beast Within or Without?

Jennifer Lin

Web technologies have dramatically changed the ways and speed in which information is exchanged and spreads across people, machines, and the social systems tying them together. PLOS, an open-access, digital-only academic publisher, sought to leverage this potential to transform scholarly communications. Readers access, share, critique, discuss, and recommend the scholarly research published online. But like other publishers, it knew almost nothing about the dissemination, reception, and impact of science and biomedical research it published when it launched its seven journals.

Often, journals cater to a particular subject area domain and thus the impact of publications may be examined locally. PLOS ONE though was the first "mega-journal" covering all scientific disciplines and thus faced a greater challenge. The scholarly community relied heavily on the Impact Factor at that point even though its flaws were widely acknowledged. It was heavily gamed, slow to accumulate, and overlooked all but citations as contributing factors (PLOS Medicine Editors, 2006; Falagas and Alexiou, 2008; Priem et al., 2010). There simply were no practical alternatives available at scale to publishers, no less to serve all scholarly literature.

And so PLOS launched its Article-Level Metrics (ALM) program in 2009. PLOS acted with a view to move from a journal-based communication system, whereby research articles are sorted into journals before publication, toward an article-based system in which articles are judged on their own merits rather than on the basis of the journal in which they are published. During my employment there, I worked on a team whose mandate was to find more effective ways to capture the diverse traces of dissemination from online activities involving PLOS's articles across the web. And we developed tools to support searching, filtering, tracking, organizing, and mining relevant to readers on the basis of these measurements. This was especially useful for helping readers navigate the largest scholarly journal, PLOS ONE. Beyond the improvement of content delivery capabilities, we also incorporated them into editorial and author services. ALMs were used to identify and target new research areas that were beginning to flourish so that we could better support scholarly communication in these emerging communities. We also used ALMs to provide authors a real-time view into what happens to their paper from the point of publication onward.

But our fuller vision required the availability of these metrics for all scholarly publications regardless of publisher. So we evangelized across the research ecosystem at the same time as developing open-source software for others to collect and display altmetrics. In so doing, we found that research funders did not know the reach and impact of the research they had supported. Institutions needed to understand their role in supporting research and did not have a systematic view of their faculty's scholarly contributions. And as altmetrics developed, they increasingly found that this emergent class of indicators could play a role in answering these questions (Dinsmore et al., 2014).

Around the same time as the birth of PLOS ALM, a group of researchers interested in the development of scholarly impact measures based on activity in online tools and environments came together under the banner of "altmetrics" and codified with their founding document, the Altmetrics Manifesto (Priem et al., 2010). Given the overlap in views and aims, PLOS's ALM work joined up with the early altmetrics efforts to catalyze change together.

The name altmetrics—coined as a collective noun for a *class of metrics*—performs a hefty job in accommodating a diverse range of online activity. It works by way of exclusion. Rather than pointing to a single indicator, it is defined against that which it is "alt-," the metric of formal literature citations (or any built off it). It incorporates online events surrounding scholarly objects (i.e., links to them) as far ranging as news media and blogging aggregators, online reference encyclopedias, social media, recommendation services, educational resource indexers, technology commercialization indexers, and reference manager and academic social network sites. Altmetrics shares the dependence of the World Wide Web with webometrics and cybermetrics, but is focused on applications for research discovery and assessment.

As an ever-evolving class of metrics, no canonical or definitive list exists. Those commonly included are article page views; downloads; comments on the publisher platform; shares on Facebook, LinkedIn, and Twitter; Zotero and CiteULike social bookmarks; Wikipedia references; mentions on Reddit or Stack Exchange discussion boards; and shares on Mendeley or other social network sites for researchers. New sites may emerge to gain popularity amongst certain scholarly communities, while others disappear or fall out of favor. Such shifts in online social behavior will be duly reflected in altmetrics. Even classificatory schemes (Lin and Fenner, 2013), which organize this buzzing basket of metrics to facilitate the study and applications of it, do not handle the rapid evolution of online research activity well.

To date, altmetrics remains a relatively new field, far from a mature one, and its reception has been quite varied. But as the data become more interesting and the utility more discernible, the emergence and explosion of these metrics has provoked and further escalated larger questions about the nature and role of research assessment, public value versus scholarly value, the economics and politics of supporting research, as well as the underlying assumptions embedded in quality and impact. It is particularly noteworthy that altmetrics has served as both gadfly and whipping post, contesting the monolithic view of what impact means as represented by the Journal Impact Factor. In fact, my PLOS team sought out ALM and altmetrics to force a larger discussion on the "qualities of quality" (or "Excellence") (Moore et al., 2017), explicitly broadening the fundamentally heterogeneous concept, which has been historically flattened due to data deficiencies and the false equating of simplicity and efficiency (Hill, 2017).

So altmetrics was born out of and continues to occupy a highly fraught space. And the debates concerning this new class of metrics too frequently lands on the susceptibility of intentional manipulation for altmetrics, here defined as gaming (Priem et al., 2012; Holmberg, 2014, 2015). In this chapter, I seek to address the anxiety of altmetrics gaming by locating and resituating its attendant issues within a broader context of data irregularities at large (inherent to all systems). I then outline a set of technical and governance needs for the research community to establish data integrity and more importantly to develop trust in this basket of new metrics.

The Anxiety of Gaming

The application of metrics may mitigate conflict, overcome distrust, and coordinate resource allocation. Yet metrics also produce unintended consequences. They become part of a reflexive sense-making dynamic, which shapes new perceptions, alters behavior as well as the narratives used to support decisions (including funding and professional advancement). These reactive and performative technologies take on the quality of self-fulfilling prophecies (Espeland and Sauder, 2007) by producing behavior changes that persist over time, altering both the expectations of those subject to them and the environmental conditions that reward particular outcomes within a specific incentive structure.

That Goodhart's law joins up metrics and gaming as an intrinsic connection (Goodhart, 1985) gives us pause in the well-studied area of scholarly metrics, specifically with citation-based forms of assessment (Franck, 1999; Opatrny, 2008; Delgado et al., 2012; Tuchman, 2012; Wilhite and Fong, 2012). And with the emergence of its alternative in altmetrics, we see this played out writ large in the debates surrounding its usefulness and validity. Altmetrics gaming—the intentional manipulation of the online activity measured—is a legitimate concern and warrants thoughtful consideration, considering the potential gravity of the offense. But these trepidations are not substantiated while basic applications in the wild have been so limited. While there have been accounts of researchers including them in CVs and bios, altmetrics do not currently play a formal, significant role in the allocation of funding resources or academic postings, where stakes are considerable, and so the anxiety surrounding altmetrics is presently best understood as speculative (Adie, 2013; Holmberg, 2014).

In fact, altmetrics may be harder to game as a suite of metrics. The technical barrier is higher for this multidimensional set of measurements compared to citations alone. Artificial citations on one or multiple papers can automatically throw the Journal Impact Factor, h-index, Eigenfactor, and other citation-based metrics for the paper(s) affected into question. However, gaming altmetrics requires manipulating measurements across a diverse set of independent web platforms. This involves extensive coordination of multiple methods specific to each metric. In addition to the lack of formal incentive, there are simply no easy ways to do it. We might see increasing sophistication in illicit tools available to coordinate online usage manipulation across the internet in the future. But currently, the anxiety of gaming may be warranted but is a threat far more conceptual than material.

An Initial Characterization of Gaming

Discussions on altmetrics gaming up to now have largely sidestepped the prior question of art: *what is it*? We have no well-accepted characterization of altmetrics gaming, especially one that accounts for the complex and continually evolving information exchanges across the web. In turn, we also have not established effective strategies that address the issue as so defined.

Gaming might be relatively simple to describe with traditional metrics. But it proves to be a real challenge for altmetrics. The view under the hood is a swarm of unceasing activity occurring all across the globe: both humans and machines interacting with online research objects. These are independently measured, collated, and then processed and fed into analytics, reports, visualizations, algorithms, and search and discovery filters. Source platforms capture activity on their systems. Raw data aggregators collate the data from data sources. Altmetrics systems select, clean, enrich, package, deliver, interpret, and present data as altmetrics. The supply chain is an expansive system of organizations, people, activities, information, and resources involved in producing and distributing the data, including upstream and downstream flows from the numerous sites where research activity occurs to any party across the research ecosystem that uses them. By and large, altmetrics providers at present moment occupy all points in the production chain.

In this distributed information environment, data irregularity may take different forms with different causes and effects. *Is this instance gaming or is it merely irregular?* We need an analytical approach that first surfaces data irregularities and then determines their nature. To echo Paul Wouters's earlier chapter, context is critical and arguably even more so in the case of manipulation allegations. We can outline the effects of gaming, and then like a detective in a police procedural, recreate the setting for the crime. Setting is context here, and context is the key to solving the crime. Irregularity can only be defined from a normative baseline of expected levels that are specific to each website ("data source"). But altmetrics are dynamic and reflect the changing tastes and whims of communities of practice as they engage with research objects. Cell biologists might adopt Mendeley or Twitter to discuss literature four years before historians, two years before they begin to use Reddit. Additionally, altmetrics activity has many dimensions. Usage varies greatly by research object based on, for example, subject area and age, and its profile is specific to each source platform. For example, Twitter activity begins and dissipates rapidly after publication compared to blog posts and reference works. The effectiveness of a baseline will depend on its ability to accommodate changing conditions and wide variability of temporality between the types of activity measured. At present moment, bibliometric scholars are only beginning to understand these patterns, which are prerequisite to establishing "regular" baselines.

In my view, data irregularities fall into four overall categories at the highest level (table 16.1) (Lin, 2012). In type 1, suspicious activity arises

Table 16.1
Typology of altmetrics data irregularities

Type 1	Suspicious activity on a particular platform that is inconsistent with previous patterns due to erroneous or outdated parameters that set normal (i.e., expected) levels of activity
Type 2	Elevated activity for publications that garner significant attention by the research community or populace at large, naturally reflected in online activity (for example, breakthrough results and novel protocol)
Type 3	Third-party activity disassociated with any express interest in the research publication and its measurements (for example, link farms, bots, and spam devices)
Type 4	Fraudulent activity caused by a party or set of parties whose intent is to willfully manipulate the measurements of online activity

on a particular platform that is inconsistent with previous patterns due to erroneous or outdated parameters that set normal (i.e., expected) levels of activity. Researcher adoption of particular online channels is prone to waxing or waning over time. Ecologists may choose to use Mendeley but switch to ResearchGate the following year. Unless the baseline is continually updated, we may register natural behavioral changes as irregular activity in the measurement observed. Type 2 includes publications that garner significant attention (for example, breakthrough results and novel protocol) by the research community or populace at large, and this is naturally reflected in the measurement of online engagement. Type 3 includes third-party activity disassociated with any express interest in the research object per se (or the measurements of the activity surrounding it). Link farms, "bots," and other online spam devices are prevalent on the web. And they have no intention of manipulating any measurements captured from their activity. Only type 4 concerns the fraudulent activity of gaming that remains the focus of our discussion: gaming via willful manipulation. It entails both act and intention.

Gaming can theoretically occur at any point in the altmetrics data chain, but it will likely occur where activity is measured (for example, on Twitter, Wikipedia, and Mendeley), prior to the processing and delivery of altmetrics where an agent might access the platform to inflate the counts. (Manipulation further downstream would require directly subverting system security to alter altmetrics calculations or representations. This is better characterized as hacking.)

To effectively address gaming, both technology capacities and community governance mechanisms need to be established. By and large, neither have, yet as core technologies for producing and provisioning altmetrics data continue to advance, interest in them is still selective, and experimentation with its applications have been early, exploratory, and often naive. But we have sufficient knowledge and experience from scholarly communications to know that any set of solutions will need to be scalable and flexible to adapt to the ever-changing nature of altmetrics and the incredible rate of growth in scholarly communications.

Altmetrics Gaming

Altmetrics needs robust technology that can scale to support the growing volume of published literature, information security controls, and systems with high availability and performance as well as data accuracy and consistency mechanisms. While we have a general characterization of altmetrics data irregularities, we need to identify it in the systems that generate the metrics. Automated monitoring and auditing is critical here. Currently, the websites or platforms where activity occurs, raw data aggregators, and the systems that compute altmetrics all employ their own approaches. Some may actively police and exclude suspicious behavior when it begins to occur. Others conduct passive monitoring and retroactively resolve issues (Gordon et al., 2015). But broadly speaking, identification entails a two-step process that first surfaces data irregularities and then ascertains that data was intentionally manipulated to gain some scholarly advantage.

Trend and event detection algorithms can hone in on our object of interest in the vast sea of data to signal the possibility of dodgy behavior. These are well established for citations and make manipulation possible to detect (McVeigh, 2002). For altmetrics, such applications will need to be paired with a robust normalization strategy that accommodates wide variations in different communities of practice by discipline and country, for example. Early studies on the altmetrics correlations have uncovered preliminary associations between metrics (Eysenbach, 2010; Priem et al., 2012; Shaui et al., 2012; Liu et al., 2013; Thelwall et al., 2013; Zahedi et al., 2014). As these findings mature into significant results, statistical experts can develop more sophisticated ways to establish baseline levels for expected counts. Cross-validation of data sources is then helpful to ascertain whether the irregularity is due to real interest in a paper where signals are registered across websites (type 2) or whether an agent (or a coordinated group) is artificially driving up counts on a single or subset of websites (type 3 or 4).

Also, pattern recognition across multiple sources may offer an even more consistent basis for detection. Early pioneers such as Scott Chamberlain from rOpenSci have begun to prototype open source tooling for gaming detection based on correlations among metrics (Chamberlain, 2015). Additionally, common statistical heuristics (e.g., Kleinberg burst analysis, hidden or semi-hidden Markov models, switching Poisson process, and Rank Surprise method) employed in other settings may be appropriate for altmetrics monitoring (PLOS, 2012). Machine learning advancements would also prove its worth in spades here. The activity distributions used to define normal behavior are hardly static, but new computational systems might offer dynamic activity profiles that automatically update as online social behaviors evolve.

Neither type 1 nor 2 constitutes occasions of manipulation, so the measurements do not need correcting. But once irregularities have been found in types 3 and 4, data cleanup may be needed so that it can be used (and more importantly trusted once more). Reprocessing activity counts, however, may be an expensive procedure in many systems. But with data management a principal feature, prudent technology design can make data adjustments a relatively simple affair. Additionally, updates to monitoring mechanisms may be necessary if the data irregularity falls in type 1 or 3, such as baseline adjustments or blocks to IP addresses that prevent future hits, respectively. We need to establish shared conventions (excluded sources, data adjustment practices) adopted by all parties involved in altmetrics data production and management. Without as much, discrepancies in the metrics data may have deleterious effects in the usefulness of altmetrics, especially when establishing trust is of paramount importance.

The technological infrastructure for altmetrics may help us identify data irregularities and instances of gaming with more research and development. But effective handling of gaming needs to go beyond the tools needed to identify and clean it up. Here, community is critical to coordinating an overarching behavioral framework of self-regulatory mechanisms supported by system incentives and sanctions. Protocols are already in place for treating academic misconduct as a professional offense and would serve as a sound basis for altmetrics gaming. Research institutions and funders have existing processes and personnel empowered to investigate and take administrative action on allegations of research misconduct such as plagiarism, mistreatment of human and animal subjects, and manipulation of results. This infrastructure need not be replicated and could be extended to include cases of altmetrics gaming. We also need clear academic norms that spell out appropriate or inappropriate behaviors. Academia's current ethos of self-policing further reinforces individual adherence to norms and offers a solid basis for effective accountability.

The scholarly incentive structure at large can either cultivate positive behaviors or instigate more prevalent cases of gaming. As the ability to detect gaming increases with the development and application of these tools, the risk of engaging in such practices would magnify, thereby lowering the incentive to engage in this behavior. But responsible use of altmetrics is the strongest beachhead for responsible production of altmetrics (Neylon, 2014). Here, the explicit appeal from Higher Education Council for Education's report for the UK Research Excellence Framework is a critical contribution to appropriate use of quantitative indicators in the governance, management, and assessment of research (Wilsdon et al., 2015).

Data Integrity: The Real Altmetrics Issue

All this may dampen the frequency of gaming, but I argue that an inverse optics is needed here. Instead of hunting down gaming offenses, we apply ourselves to creating an environment most conducive to overall altmetrics data integrity. In the networked world of the scientific enterprise, data integrity is a shared responsibility of all the players involved. And in the distributed network of the altmetrics ecosystem, this proves to be just as true. This environment is made up of multiple agents that capture activity originating on their site, data aggregators who collect the data, and distributors who enrich and package the data along with any host of additional intermediaries. To understand data integrity from a network standpoint, we recognize the diversity of players and complex information exchanges across the web that occur at each and every site. Key parties not only include the aggregators and altmetrics systems, but also the source platforms where activity occurs. This also includes all consumers of altmetrics data as well: funders and research institutions as well as the technology services that act as intermediaries.

Just as scholarly infrastructure underlies the operations of the research enterprise, altmetrics infrastructure needs to be a principal part of ensuring oversight and trust in the health of altmetrics. The latest intervention in altmetrics gaming is the establishment of a central archive of raw data from which altmetrics can be created. The bedrock for altmetrics infrastructure is the provisioning and preservation of underlying data generated by the actions of the research community, which becomes the basis of calculation of the metrics. Here, the events from online platforms are best treated as a common resource, so the community can use them to inform decisions as equally as private enterprise can develop services powered by the data (Bilder et al., 2016). And of importance here, the community can address data irregularities and identify instances of gaming. The National Information Standards Organization (NISO) has taken the helm in leading community discussions on setting standards and best practices for the development and collection of altmetrics in their Alternative Metrics Initiative. There is agreement here—NISO calls out a centralized data clearinghouse as a key requirement in the antigaming recommendations (NISO, 2014).

As such, I now work with Crossref, a scholarly infrastructure organization, to fill this need. Our Event Data service collects underlying data for online activity surrounding publications across the web at scale (i.e., to encompass all published literature) and makes it freely available to all. This piece of altmetrics infrastructure effectively detaches upstream event aggregation from downstream altmetrics services in the altmetrics supply chain, making specialization possible in the production process and increasing efficiency in the entire system. It also largely resolves the current lack of data standardization (including definition of a common baseline, exchange mechanism employed, and construction of the queries) that poses a particular challenge to gaming detection. Furthermore, this structural shift alleviates the altmetrics gaming problem by localizing it to specific, discrete areas of the supply chain: either the data source platform or the altmetrics providers where altmetrics are computed from the raw event activity distributed from the central Crossref Event Data archive.

In addition to the underlying data, data integrity needs community-based standards that ensure consistent aggregation of altmetrics as well as transparency measures that could serve as a solid basis for data reliability and trustworthiness. While players need not necessarily adopt a single process or technology for data to be trusted, the black box nature of operations leads to a significant degree of unknowability and thus uncertainty. NISO asks aggregators and altmetrics providers to report on how they collate events and calculate their measures in their Altmetrics Data Quality Code of Conduct. They also ask altmetrics providers to describe how they have kept their data free of error (NISO, 2016). Disclosures support public accountability and are beachheads for the detection and resolution of altmetrics gaming.

Regular data audits by a trusted, independent party can also serve as additional transparency measures and create a more resilient environment against gaming. With citations, underlying data is theoretically available in persistent and consistent form (even if not openly available). This is not currently true with altmetrics, where data is currently ephemeral and opaque. But scholarly infrastructure would make it possible for a third party to conduct audits on the archive of data, rather than at the site of each altmetrics provider. If data are made openly available to all, audits can be conducted not only by dedicated parties entrusted by the community, but also by any member in the community. These provide yet additional layers of support for data integrity overall as well as insurance against altmetrics gaming. In addition, other supporting measures have been proposed to bolster altmetrics data integrity, including a public, open reference dataset for proper metrics development and auditing as well as open-source analytical tools used on the dataset for true transparency and reproducibility of the metrics (Lin et al., 2017).

Concluding Reflections

I foresee that altmetrics, this latest intervention in the scholarly research enterprise, will likely play some role in the future of research. The activity of sharing, discussing, and critiquing ideas is fundamental to the progress of research, and these modes of interaction are already prevalent online. Considering the uptake (and potential gains) so far in capturing and tracking these activities as a part of, for example, driving research management and literature discovery, altmetrics is beginning to offer some advantages. In what ways we will capture its value remains the big unanswered question, however. And whether they will be conducive to a healthy research environment is yet another. These rest on many practical factors, including availability and reliability of data (i.e., public provisioning), bibliometric understanding of what the data "means" based on the nature of activity involved, and meaningful use cases supported by tools/systems/platforms across communities in the research ecosystem. To establish trust in any application of altmetrics, the question of gaming needs to be considered in context and take an environmental approach that includes supporting technology and standard practices and norms, as well as community engagement across research institutions and funders. But as gaming is inherent in any metrics system in the same way that data integrity—valid and reliable data—is an issue inherent in all information systems, the broader view of requirements for data quality and integrity is also critical. With adequate community infrastructure to support broad development and appropriate use of altmetrics, the problem and anxiety of gaming may end up much less a beast within and rather one without.

Notes

Acknowledgments: I thank Cameron Neylon, Geoffrey Bilder, Joe Wass, Alexandra Lippman, and Mario Biagioli for comments that greatly improved the chapter.

References

Adie, Euan. 2013. “Gaming Altmetrics.” *Altmetric Blog*. Accessed May 29, 2016. *http://www.altmetric.com/blog/gaming-altmetrics/*.

Bilder, Geoffrey, Jennifer Lin, and Cameron Neylon. 2015. “Principles for Open Scholarly Infrastructures-v1.” figshare. Accessed June 10, 2017. *https://doi.org/10.6084/m9.figshare.1314859.v1*.

Bilder, Geoffrey, Jennifer Lin, and Cameron Neylon. 2016. “Where Are the Pipes? Building Foundational Infrastructures for Future Services.” *Science in the Open*. Accessed June 1, 2016. *http://cameronneylon.net/blog/where-are-the-pipes-building-foundational-infrastructures-for-future-services/*.

Chamberlain, Scott. 2015. “Summary of PLOS ALM Data Quality Work.” *PLOS Tech Blog,* January 14. Accessed May 29, 2016. *http://blogs.plos.org/tech/summary-of-alm-data-quality-work/*.

Costas, Rodrigo, Zohreh Zahedi, and Paul Wouters. 2014. “Do ‘Altmetrics’ Correlate with Citations? Extensive Comparison of Altmetric Indicators with Citations from a Multidisciplinary Perspective.” *Journal of the Association for Information Science and Technology*. 66:2003–2019. *doi:10.1002/asi.23309*.

Delgado Lopez-Cozar, Emilio, Nicolas Robinson-Garcia, and Daniel Torres Salinas. 2012. “Manipulating Google Scholar Citations and Google Scholar Metrics: Simple, Easy, and Tempting.” *EC3 Working Papers* 6:29. May 2012. *http://arxiv.org/abs/1212.0638*.

Dinsmore, Adam, Liz Allen, and Kevin Dolby. 2014. “Alternative Perspectives on Impact: The Potential of ALMs and Altmetrics to Inform Funders About Research Impact.” *PLOS Biol* 12(11):e1002003. *https://doi.org/10.1371/journal.pbio.1002003*. *http://arxiv.org/abs/1212.0638*.

Espeland, Wendy Nelson, and Michael Sauder. 2007. “Rankings and Reactivity: How Public Measures Recreate Social Worlds.” *American Journal of Sociology* 113(1):1–40. *https://doi.org/10.1086/517897*.

Eysenbach, Gunther. 2011. “Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact.” *J Med Internet Res*.13(4):e123. *https://doi.org/10.2196/jmir.2012*.

Falagas, Matthew, and Vangelis Alexiou. 2008. "The Top-Ten in Journal Impact Factor Manipulation." *Archivum Immunologiae et Therapiae Experimentalis* 56:223–226. *https://doi.org/10.1007/s00005-008-0024-5*.

Franck, Georg. 1999. "Scientific Communication—A Vanity Fair?" *Science* 286:53–55. *https://doi.org/10.1126/science.286.5437.53*.

Goodhart, Charles. 1981. "Problems of Monetary Management: The U.K. Experience." Anthony S. Courakis (ed.), *Inflation, Depression, and Economic Policy in the West*. (Rowman & Littlefield): 111–146. *https://doi.org/10.1007/978-1-349-17295-5_4*.

Gregg, Gordon, Jennifer Lin, Richard Cave, and Ralph Dandrea. 2015. "The Question of Data Integrity in Article-Level Metrics." *PLOS Biology* 13:e1002161. *https://doi.org/10.1371/journal.pbio.1002161*.

Hill, Steven, 2015. Higher Education Funding Council blog. Published July 27, 2015. Accessed June 10, 2017. *http://blog.hefce.ac.uk/2015/07/27/dimensions-of-quality-research/*.

Holmberg, Kim. 2014. "How Concerned Should We Be About Gaming Altmetrics?" *Social Science Research blog*, July 11. Accessed May 30, 2016. *http://www.socialsciencespace.com/2014/07/altmetrics-may-be-able-to-help-in-evaluating-societal-reach-but-research-significance-must-be-peer-reviewed/*.

Holmberg, Kim. 2015. Altmetrics for Information Professionals. *http://www.sciencedirect.com/science/book/9780081002735*.

Lin, Jennifer. 2012. "A Case Study in Anti-Gaming Mechanisms for Altmetrics: PLoS ALMs and DataTrust." *ACM Web Science Conference 2012 Workshop*. Evanston, IL. Accessed May 29, 2016. *http://altmetrics.org/altmetrics12/lin/*.

Lin, Jennifer, Fiona Murphy, Michael Taylor, and Liz Allen. 2017. "Building the Infrastructure to Make Science Metrics More Scientific." *F1000Research* 5:2897. *https://doi.org/10.12688/f1000research.10422.2*. *http://blog.crossref.org/2015/09/det-poised-for-launch.html*.

Lin, Jennifer, and Martin Fenner. 2013. "Altmetrics in Evolution: Defining & Redefining the Ontology of Article-Level Metrics." *Information Standards Quarterly* 25(2):20–26. *https://doi.org//10.3789/isqv25no2.2013.04*.

Liu, Chun Li, Yue Quan Xu, Hui Wu, Si Si Chen, and Ji Jun Guo. 2013. "Correlation and Interaction Visualization of Altmetric Indicators Extracted from Scholarly Social Network Activities: Dimensions and Structure." *Journal of Medical Internet Research* 15(11):e259. *https://doi.org/10.2196/jmir.2707*.

McVeigh, Marie. 2002. "Journal Self-Citation in the Journal Citation Reports." *Thompson Reuters White Papers*. Accessed June 1, 2016. *http://wokinfo.com/essays/journal-self-citation-jcr/*.

Moore, Samuel, Cameron Neylon, Martin Paul Eve, Daniel Paul O'Donnell, and Damian Pattinson. 2017. "'Excellence R Us': University Research and the Fetishisation of Excellence." *Palgrave Communications* 3, Article number: 16105. *https://doi.org/0.1057/palcomms.2016.105*. *http://wokinfo.com/essays/journal-self-citation-jcr/*.

Neylon, Cameron. 2014. "PLOS Submission to HEFCE RFI on Metrics in Research Assessment." *figshare*. *https://doi.org/10.6084/m9.figshare.1089555.v1*. *https://dx.doi.org/10.6084/m9.figshare.1089555.v1*.

NISO. *RP-25–201x-3, Altmetrics Data Quality Code of Conduct—Draft for Public Comment*. Accessed June 19, 2017. *http://www.niso.org/apps/group_public/document.php?document_id=16121&wg_abbrev=altmetrics-quality*.

NISO. 2014. *NISO Altmetrics Initiative Phase 1 White Paper*. Accessed June 19, 2017. *http://www.niso.org/apps/group_public/document.php?document_id=13809&wg_abbrev=altmetrics*.

Opatrny, Tomas. 2008. "Playing the System to Give Low-Impact Journal More Clout." *Nature* 455:7210. *https://doi.org/10.1038/455167b*.

PLOS. 2012. "ALM Workshop 2012 Report." *Figshare*. *https://doi.org/10.6084/m9.figshare.98828.v2*.

PLOS Medicine Editors. 2006. "The Impact Factor Game." *PLOS Med* 3(6):e291. *https://doi.org/10.1371/journal.pmed.0030291*.

Power, Michael, Tobias Scheytt, Kim Soin, and Kerstin Sahlin. 2009. "Reputational Risk as a Logic of Organizing in Late Modernity." *Organization Studies* 30:301–324. *https://doi.org/10.1177/0170840608101482*.

Priem J., P. Groth, and D. Taraborelli. 2012. "The Altmetrics Collection." *PLOS One* 7(11):e48753. *https://doi.org/10.1371/journal.pone.0048753*.

Priem, J., D.Taraborelli, P. Groth, and C. Neylon. 2010. *Altmetrics: A Manifesto*. Accessed May 28, 2016. *http://altmetrics.org/manifesto*.

Priem, Jason, Heather Piwowar, and Brad Hemminger. 2012. "Altmetrics in the Wild: An Exploratory Study of Impact Metrics Based on Social Media." *arXiv*. arXiv:1203.4745.

Shuai, X., A. Pepe, and J. Bollen. 2012. "How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations." *PLOS One* 7(11):e47523. *https://doi.org/10.1371/journal.pone.0047523*.

Strathern, Marilyn. 2000. *Audit Culture*. New York: Routledge.

Thelwall, Mike, Stephanie Haustein, Vincent Larivière, and Cassidy Sugimoto. 2013. "Do Altmetrics Work? Twitter and Ten Other Social Web Services." *PLOS One* 8(5):e64841. *https://doi.org/10.1371/journal.pone.0064841*.

Tuchman, Gaye. 2012. "Commodifying the Academic Self." *Inside Higher Ed*, February 6. Accessed May 30, 2016. *https://www.insidehighered.com/views/2012/02/06/essay-gaming-citation-index-measures*.

van Noorden, Richard. 2013. "Brazilian Citation Scheme Outed." *Nature News* 500:510–511. *https://doi.org/10.1038/500510a*.

Wilhite, A., and E. Fong. 2012. "Coercive Citation in Academic Publishing." *Science* 335(6068):532–543. *https://doi.org/10.1126/science.1212540*.

Wilsdon, James, Liz Allen, Eleonora Belfiore, Philip Campbell, Stephen Curry, Seven Hill, Richard Jones, Roger Kain, Simon Kerridge, Mike Thelwall, Jane

Tinkler, Ian Viney, Paul Wouters, Jude Hill, and Ben Johnson. 2015. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. https://doi.org/10.13140/RG.2.1.4929.1363*.

Zahedi, Zohreh, Rodrigo Costas, and Paul Wouters. 2014. "How Well Developed Are Altmetrics? A Cross-Disciplinary Analysis of the Presence of 'Alternative Metrics' in Scientific Publications." *Scientometrics* 101:1491–1513. *https://doi.org/10.1007/s11192-014-1264-0*.